Journal of Econometrics 3 (1975) 387–390. © North-Holland Publishing Company

A RESULT ON THE SIGN OF RESTRICTED LEAST-SQUARES ESTIMATES*

Edward E. LEAMER

Harvard University, Cambridge, Mass. 02138, U.S.A.

Received December 1975

When a variable is dropped from a least-squares regression equation, there can be no change in sign of any coefficient that is more significant than the coefficient of the omitted variable. More generally, a constrained least squares estimate of a parameter β_j must lie in the interval $(\hat{\beta}_j - V_{jj}^{\frac{1}{2}}|t|, \hat{\beta}_j + V_{jj}^{\frac{1}{2}}|t|)$ where $\hat{\beta}_j$ is the unconstrained estimate of β_j, $V_{jj}^{\frac{1}{2}}$ is the standard error of $\hat{\beta}_j$ and t is the t-value for testing the (univariate) restriction.

Least-squares estimates of the coefficients of a linear regression model often have signs that are regarded by the researcher to be 'wrong'. In an effort hopefully to obtain 'right' signs, statistically insignificant variables are sometimes dropped from the equation. Surprisingly enough, there can be no change in sign of any coefficient that is more significant than the coefficient of the omitted variable. More generally, a constrained least-squares estimate of a parameter β_j must lie in the interval

$$(\hat{\beta}_j - V_{jj}^{\frac{1}{2}}|t|, \hat{\beta}_j + V_{jj}^{\frac{1}{2}}|t|),$$

where $\hat{\beta}_j$ is the unconstrained estimate of β_j, $V_{jj}^{\frac{1}{2}}$ is the standard error of $\hat{\beta}_j$ and t is the t-value for testing the (univariate) restriction.

A proof of this proposition is straightforward. The regression process may be written in matrix form as

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{u}, \tag{1}$$

where $\boldsymbol{Y}$ is a $T \times 1$ observable vector, $\boldsymbol{X}$ is a $T \times k$ observable matrix, $\boldsymbol{\beta}$ is a $k \times 1$ unobservable vector, and $\boldsymbol{u}$ is a $T \times 1$ unobservable vector. Let the least-squares estimate of $\boldsymbol{\beta}$ be

$$\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'Y, \tag{2}$$

*Support from NSF grant GS 31929 is gratefully acknowledged. Comments by Yair Mundlak on an earlier version led to the more general results presented here.

the variance–covariance matrix be[1]

$$V = (X'X)^{-1},$$

and a t-statistic for testing $\beta_i = 0$ be

$$t_i = \hat{\beta}_i / V_{ii}^{1/2}.$$

The least-squares estimate of $\boldsymbol{\beta}$ constrained to $\boldsymbol{R\beta} = \boldsymbol{r}$ is[2]

$$\hat{\boldsymbol{\beta}}^* = \hat{\boldsymbol{\beta}} - \boldsymbol{V R'}[\boldsymbol{R V R'}]^{-1}[\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r}]. \tag{3}$$

We shall consider constraining β_i to r in which case $\boldsymbol{R}$ is a row vector with a one in the ith column and zeroes elsewhere, and we may write (3) as

$$\hat{\boldsymbol{\beta}}^* = \hat{\boldsymbol{\beta}} - \boldsymbol{V}_i(\hat{\beta}_i - r)/V_{ii}, \tag{4}$$

where $\boldsymbol{V}_i$ is the ith column (or row) of $\boldsymbol{V}$.

For expository purposes, we will prove the least general result first.

Lemma. The least-squares estimate of β_j and the constrained least-squares estimate of β_j, with β_i set to zero, have the same sign if the t-statistic of β_i is less in absolute value than the t-statistic of β_j.

Proof. We wish to show that, if $t_i^2 \leqq t_j^2$, then $\hat{\beta}_j$ and $\hat{\beta}_j^*$ have the same sign. Using (4) we can write

$$\hat{\beta}_j^* = \hat{\beta}_j - V_{ji}\hat{\beta}_i / V_{ii} = V_{jj}^{1/2}[t_j - V_{ji}t_i/(V_{ii}V_{jj})^{1/2}]. \tag{5}$$

The positive definiteness of $\boldsymbol{V}$ implies that

$$-1 \leqq V_{ji}/(V_{ii}V_{jj})^{1/2} \leqq 1.$$

Using this fact in eq. (5) together with the hypothesis $t_i^2 \leqq t_j^2$, makes the result obvious. More formally, $\hat{\beta}_j$ and $\hat{\beta}_j^*$ have the same sign if and only if

$$0 \leqq \hat{\beta}_j^* \hat{\beta}_j = V_{jj}[t_j^2 - V_{ji}t_i t_j/(V_{ii}V_{jj})^{1/2}].$$

But by hypothesis $|t_j| > |t_i|$, and

$$t_j^2 \geqq |t_i t_j| \geqq |V_{ji}t_i t_j/(V_{ii}V_{jj})^{1/2}| \geqq V_{ji}t_i t_j/(V_{ii}V_{jj})^{1/2}.$$

[1] We have set the process variance to one without affecting the result below.
[2] See for example, Goldberger (1964, p. 257).

Incidentally, this result, for the special case of dropping the variable with the smallest t, is geometrically 'obvious'. It amounts to the observation that the vector 'closest' to the least-squares vector but lying in one of the hyperplanes, $\beta_i = 0$, must be in the same orthant as the least-squares vector. Suppose it were not and suppose that $\hat{\boldsymbol{\beta}}^*$ were such a vector computed by least-squares subject to the constraint $\beta_i = 0$. Then the ellipsoid $(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'X'X(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}) = Z$ tangent to the hyperplane $\beta_i = 0$ at the point $\hat{\boldsymbol{\beta}}^*$ could not intersect the boundary of the orthant of $\hat{\boldsymbol{\beta}}$ since, if it did, there would be some other variable with a lower t-value. But every ellipsoid must travel into the orthant of $\hat{\boldsymbol{\beta}}$.

There is nothing special about the locations of the hypotheses in this result and we leave to the reader the verification of the following corollary which bounds the constrained estimate on one side of some arbitrary point, c.

Corollary 1. The least-squares estimate of $\beta_j - c$ and the constrained least-squares estimate of $\beta_j - c$, with β_i set to r, have the same sign if the t-statistic for testing $\beta_i = r$ is less in absolute value than the t-statistic for testing $\beta_j = c$.

For any constrained point r in Corollary 1 we can solve for the set of points c, all of which bound the constrained estimate. The intersection of these bounds implies the following most general univariate result.

Theorem 1. The constrained least-squares estimate of β_j must lie in the interval $(\hat{\beta}_j - V_{jj}^{1/2}|t|,\ \hat{\beta}_j + V_{jj}^{1/2}|t|)$ where t is the t-statistic for testing the univariate restriction.

The multivariate generalization of this bound is less useful since it involves F-statistics which are computed with greater difficulty than the t-statistics discussed above.

Theorem 2. The least-squares estimate of β_j, $\hat{\beta}_j$, and the constrained least-squares estimate of β_j, $\hat{\beta}_j^*$, with the constraint $\boldsymbol{R\beta} = \boldsymbol{r}$, satisfy the inequality

$$|\hat{\beta}_j - \hat{\beta}_j^*| \leqq V_{jj}^{1/2} q^{1/2} F^{1/2},$$

where q is the rank of $\boldsymbol{R}$, F is the F-statistic for testing the constraint, and $V_{jj}^{1/2}$ is the estimated standard error of $\hat{\beta}_j$.

Proof. The likelihood ellipsoid $(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'X'X(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})$ evaluated at the constrained least-squares estimate $\boldsymbol{\beta}^*$ (3) takes on the value

$$s^2 = (\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r})'[\boldsymbol{R}V\boldsymbol{R}']^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{r}).$$

The estimate, $\boldsymbol{\beta}^*$, lies on the ellipsoid

$$(\hat{\boldsymbol{\beta}}^* - \hat{\boldsymbol{\beta}})'X'X(\hat{\boldsymbol{\beta}}^* - \hat{\boldsymbol{\beta}}) = s^2.$$

Projecting this ellipsoid onto the jth axis, Scheffe (1959, p. 410) yields the interval

$$|\beta_j^* - \hat{\beta}_j| \leqq V_{jj}^{\frac{1}{2}} s.$$

We observe in conclusion that these results can also obviously be used to explore the ordering of coefficients by considering contrasts of the form $\beta_j - \beta_k$.

References

Goldberger, A.S., 1964, Econometric theory (Wiley, New York).
Scheffe, H., 1959, The analysis of variance (Wiley, New York).